

AD-A089 763

PITTSBURGH UNIV PA INST FOR STATISTICS AND APPLICATIONS F/6 12/1
A MULTIVARIATE CORRELATION RATIO.(U)

MAY 80 A R SAMPSON

F49620-79-C-0161

UNCLASSIFIED

TR-80-6

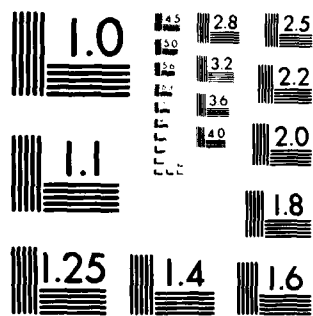
AFOSR-TR-80-0938

NL

1 1 1
AD
A089 763



END
DATE
FILMED
11 80
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

AFOSR-TR- 80-0988

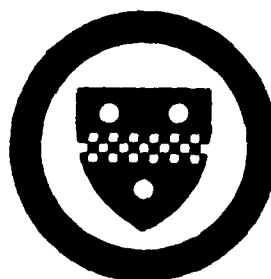
LEVEL II

4

AD A089763

DTIC
ELECTE
SEP 30 1980
S E

Institute for Statistics and Applications
University of Pittsburgh



Approved for public release;
distribution unlimited.

DOC FILE COPY

F-49620-79-C-0161

Approved for public release;
distribution unlimited.

80 9 25 122

LEVEL II

A MULTIVARIATE CORRELATION RATIO

Allan R. Sampson^{*}

Department of Mathematics and Statistics
University of Pittsburgh

May 1980

Technical Report No. 80-6

Institute for Statistics and Applications
Department of Mathematics and Statistics
University of Pittsburgh
Pittsburgh, PA 15260

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DDC
This technical report has been reviewed and is
approved for public release IAW AFR 190-12 (7b).
Distribution is unlimited.
A. D. BLOSE
Technical Information Officer

^{*}The work of this author is sponsored by the Air Force Office of Scientific Research, Air Force Systems Command under Contract F49620-79-C-0161. Reproduction in whole or in part is permitted for any purposes of the United States Government.

DTIC
ELECTE
SEP 30 1980
S D
E

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER 18 AFOSR TR-80-0938 | 2. GOVT ACCESSION NO. AD-A089 763 | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle) A MULTIVARIATE CORRELATION RATIO | 5. TYPE OF REPORT & PERIOD COVERED Interim Rept. | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s) Allan R. Sampson | 8. CONTRACT OR GRANT NUMBER(s) F49620-79-C-0161 | |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Pittsburgh Department of Mathematics & Statistics Pittsburgh, PA 15260 | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBER 61102F 2304/AS | |
| 11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332 | 12. REPORT DATE May 1980 | |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 1226 | 13. NUMBER OF PAGES 21 | 15. SECURITY CLASS. (of this report) UNCLASSIFIED |
| 16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited. 14 TR-80-6 | | |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) | | |
| 18. SUPPLEMENTARY NOTES | | |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Correlation ratio, multivariate correlation ratio, vector correlation, canonical correlation, sup-correlation, elliptically symmetric. | | |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number) A brief review of the historical background and certain known results concerning the univariate correlation ratio are given. A multivariate correlation ratio of a random vector Y upon a random vector X is defined by $\eta_A(Y;X) = \text{tr}(A^{-1} \text{Cov } E(Y X))^{1/2} (\text{tr}(A^{-1} E_Y))^{-1/2},$ | | |

UNCLASSIFIED

FTA sub A

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

where A is a given positive definite matrix. The properties of η_A are discussed, with particular attention paid to a 'correlation-maximizing' property. A number of examples illustrating the application of η_A are given; these examples include the multivariate normal, the elliptically symmetric distributions, the Farlie-Morgenstern-Gumbel family, and the multinomial. The problem of maximizing $\eta_A(BY; X)$ over suitable matrices B is considered and the results that are obtained are related to canonical correlations for the multivariate normal.

| | |
|--------------------|-------------------------------------|
| Accession For | |
| NTIS GRA&I | <input checked="" type="checkbox"/> |
| DDC TAB | <input type="checkbox"/> |
| Unannounced | <input type="checkbox"/> |
| Justification | |
| By _____ | |
| Distribution/ | |
| Availability Codes | |
| Dist. | Avail and/or special |
| A | |

UNCLASSIFIED

A Multivariate Correlation Ratio

by

Allan R. Sampson^{*}

Department of Mathematics and Statistics
University of Pittsburgh
Pittsburgh, PA 15260

ABSTRACT

A brief review of the historical background and certain known results concerning the univariate correlation ratio are given. A multivariate correlation ratio of a random vector \underline{Y} upon a random vector \underline{X} is defined by

$$\eta_A(\underline{Y};\underline{X}) = \{\text{tr}(\underline{A}^{-1} \text{Cov } E(\underline{Y}|\underline{X}))\}^{\frac{1}{2}} \{\text{tr}(\underline{A}^{-1} \underline{\Sigma}_Y)\}^{-\frac{1}{2}},$$

where \underline{A} is a given positive definite matrix. The properties of η_A are discussed, with particular attention paid to a "correlation-maximizing" property. A number of examples illustrating the application of η_A are given; these examples include the multivariate normal, the elliptically symmetric distributions, the Farlie-Morgenstern-Gumbel family, and the multinomial. The problem of maximizing $\eta_A(\underline{B}\underline{Y};\underline{X})$ over suitable matrices \underline{B} is considered and the results that are obtained are related to canonical correlations for the multivariate normal.

AMS Classification: Primary 62H20, Secondary 62J05

Key Words: Correlation ratio, multivariate correlation ratio, vector correlation, canonical correlation, sup-correlation, elliptically symmetric.

^{*}The work of this author is sponsored by the Air Force Office of Scientific Research under Contract F49620-79-C-0161. Reproduction in whole or in part is permitted for any purposes of the United States Government.

A Multivariate Correlation Ratio

by

Allan R. Sampson*

Department of Mathematics and Statistics
University of Pittsburgh
Pittsburgh, PA 15260

1. INTRODUCTION AND HISTORICAL BACKGROUND.

The correlation ratio, $\eta(Y;X)$, of a random variable Y upon a random variable X , defined by (with suitable assumptions)

$$\eta(Y;X) = \{\text{Var}(E(Y|X))/\text{Var } Y\}^{1/2}, \quad (1.1)$$

was first introduced by K. Pearson in (1903, p. 304), who wrote " η is a useful constant which ought always to be given for non-linear systems...it measures the approach of the system not only to linearity but to single valued relationship, i.e., to a causal nexus". Pearson further discussed η in his papers of (1905; or see (1948, pp. 477-528)) and (1909). In his 1905 paper, he wrote "the correlation ratio...is an excellent measure of the stringency of correlation always lying numerically between the values 0 and 1, which mark absolute independence and complete causation respectively". He further noted, based on his considerations of non-normal bivariate data, that "the ease with which η can be calculated suggests that in many cases it should accompany, if not replace the determination of the correlation coefficient".

Blakeman (1905) also introduced a criteria based on η to test for linearity of regression. Fisher (1925; pp. 257-260 of 14th Ed. (1970)),

*The work of this author is sponsored by the Air Force Office of Scientific Research under Contract F49620-79-C-0161. Reproduction in whole or in part is permitted for any purposes of the United States Government.

seemingly less enthusiastic about η , wrote concerning the sample analogue of η in the regression model that "as a descriptive statistic the utility of the correlation ratio is extremely limited". It appears that much of his concerns were based on certain distributional properties. A more recent discussion concerning η can be found in Lancaster (1969, pp. 201-202).

Various other properties of η have been considered within the literature which focuses on measures of association and measures of dependence, most of this literature having been written within the last approximately 20 years. Kruskal (1958) in his survey on ordinal measures of association discussed η , and Rényi (1959) in his axiomatic development of measures of dependence examined properties of η . More recently, Hall (1970) defined the dependence characteristic function as $\eta(e^{itY}; X)$, where a suitable extension of η to complex-valued random variables was given; the relative merits of the dependence characteristic function versus the correlation ratio were considered. Kotz and Soong (1977) further reviewed some of the probabilistic properties of η . Hall (1970) also noted that when X is vector valued; the correlation ratio $\eta(Y; X)$, defined by (1.1) now with $E(Y|X)$ in the numerator, has essentially the same properties as when X is a scalar random variable. Within a specific multivariate normal setting, Johnson and Kotz (1972, p. 186) noted that a certain multivariate beta random variable could be viewed as a multivariate generalization of η .

The correlation ratio is in some ways connected to the sup-correlation coefficient between random variables X and Y , defined by

$$\rho'(X,Y) = \sup \rho(f(X),g(Y)), \quad (1.2)$$

where the supremum is over suitable functions f and g , and ρ is the Pearson correlation coefficient. This measure of dependence was introduced by Gebelein (1941) and developed further by Sarmanov (1958A), (1958B), Rényi (1959) and Lancaster (whose work is summarized in Lancaster (1969)). Sarmanov and Zaharov (1960) extended this concept to the multivariate case, defining $\rho'(X,Y) = \sup \rho(f(X),g(Y))$, where the supremum is over suitable f,g , which map R^p and R^q , respectively, into R^1 . We note that except in very special cases, it is difficult to obtain an explicit evaluation of $\rho'(X,Y)$.

In this paper, we consider defining the correlation ratio for the case when both Y and X are vector random variables. This extension would, for example, accomodate the situation when we are studying the relationship of X_{r+1}, \dots, X_{r+s} to X_1, \dots, X_r , or when we are relating jointly a time series Y_1, \dots, Y_t to a time series X_1, \dots, X_s . The properties of this multivariate correlation ratio are explored in light of the properties of $\eta(Y;X)$. We also examine maximizing the multivariate correlation ratio over certain linear combinations, and study the relationship of this concept to other multivariate notions, including the sup-correlation. A number of specific multivariate distributional examples are considered including the normal, elliptically symmetric and Farlie-Morgenstern-Gumbel.

2. A REVIEW OF RESULTS PERTAINING TO $\eta(Y;X)$.

In this section, we survey some results concerning $\eta(Y;X)$ and discuss briefly some of the implications of these results.

Theorem 2.1. Let Y be a random variable with $0 < \text{Var } Y < \infty$.

Let \underline{X} be a p -dimensional random vector, jointly distributed with Y .

(a) Then $\min E(Y - g(\underline{X}))^2$ occurs at $g(\underline{x}) = E(Y | \underline{X} = \underline{x})$, where the minimum is over all measurable $g: R^p \rightarrow R^1$, for which $E(Y - g(\underline{X}))^2 < \infty$.

(b) Furthermore, the minimum value is $(1 - \eta^2(Y; \underline{X})) \text{Var } Y$.

A proof of essentially Theorem 2.1 can be found, for example, in Parzen (1960) or Hall (1970).

Theorem 2.2. Let Y be a random variable with $0 < \text{Var } Y < \infty$.

Let \underline{X} be a p -dimensional random vector jointly distributed with Y .

(a) Then $\max |\rho(Y, g(\underline{X}))|$ occurs at $g(\underline{x}) = E(Y | \underline{X} = \underline{x})$, where the maximum is over all measurable functions $g: R^p \rightarrow R^1$ for which the correlation is defined. (b) Furthermore, the maximum value is $\eta(Y; \underline{X})$.

Note that in Theorem 2.2 if $g(\underline{x})$ maximizes, then $\alpha g(\underline{x}) + \beta$, $\alpha \neq 0$, maximizes. A proof of Theorem 2.2 in the case X is a scalar can be found in Kotz and Soong (1977); the proof when \underline{X} is a vector is identical. This very interesting interpretation of $\eta(Y; \underline{X})$, according to Kruskal (1958), was first noted by Fréchet (1933), (1934). Earlier, Pearson (1905) had proved $\eta(Y; X) \geq |\rho(Y, X)|$ and Fisher (1925) had shown $\eta(Y; X) \geq \max |\rho(Y, g(X))|$.

An immediate result of Theorem 2.2 is that $0 \leq \eta \leq 1$. From Theorem 2.1, we observe that Y being predicted by $g(\underline{X})$ with an expected squared error of zero is equivalent to $\eta(Y; \underline{X}) = 1$. A further consequence of Theorem 2.2 is that $\eta(Y; \underline{X}) = 0$ is equivalent to $\rho(Y, h(\underline{X})) = 0$ for all measurable functions h , with $0 < \text{Var } h(\underline{X}) < \infty$. This also implies that $E(Y | \underline{X} = \underline{x}) = E(Y)$ a.e. for all \underline{x} .

One commonly perceived deficiency of η as a measure of dependence

(e.g., Renyi (1959)) is that $\eta = 0$ does not imply Y and \underline{X} are independent. The correlation ratio being zero may be interpreted as being "between" independence and uncorrelatedness (in terms of multiple correlation) in the following sense. Independence of Y and \underline{X} is equivalent to $\rho(f(Y), g(\underline{X})) = 0$ for all suitable f, g ; and uncorrelatedness is equivalent to $\rho(\alpha_1 Y + \beta_1, \alpha_2' \underline{X} + \beta_2) = 0$ for all $\alpha_1, \beta_1, \alpha_2, \beta_2$; while zero correlation ratio, as noted, is equivalent to $\rho(\alpha_1 Y + \beta_1, g(\underline{X})) = 0$, for all α_1, β_1 , and suitable g .

When $(Y, \underline{X})'$ have a joint multivariate normal distribution, the multiple correlation is defined (e.g., Anderson (1958)) by $\max_{\alpha} \rho(Y, \alpha' \underline{X})$. For the multivariate normal $E(Y|\underline{X})$ is linear in \underline{X} , so that it follows from Theorems 2.1 and 2.2 that $\eta(Y; \underline{X})$ is the same as the multiple correlation of Y with \underline{X} . For the linear prediction problem where g in Theorems 2.1 and 2.2 is restricted to be linear, it is more appropriate (e.g., Parzen (1960)) to use $\max_{\alpha} \rho(Y, \alpha' \underline{X})$ than $\eta(Y; \underline{X})$ for measuring dependence. This suggests, perhaps, that when \underline{X} is an arbitrary random vector, one might adopt the nomenclature, linear multiple correlation generalized multiple correlation, and sup-multiple-correlation, respectively for $\max_{\alpha} \rho(Y, \alpha' \underline{X})$, $\eta(Y; \underline{X})$, and $\sup_{f, g} \rho(f(Y), g(\underline{X}))$. For the multivariate normal these three measures are equal.

3. A MULTIVARIATE CORRELATION RATIO.

In this section, we consider the definition of a correlation ratio of Y upon \underline{X} , where Y is a $q \times 1$ random vector and \underline{X} is a $p \times 1$ random vector. Even though Y is now a vector, it seems suitable to have the multivariate correlation ratio take scalar values, rather than

attempt a multiple-valued version. To motivate our definition of the multivariate correlation ratio we first consider the multivariate prediction problem and proceed from there.

In referring to the covariance matrix of a random vector \underline{W} , we use the notation $\text{Cov}(\underline{W})$ and $\underline{\Sigma}_W$ interchangeably. The cross-covariance between random vectors $\underline{S}, \underline{T}$ is denoted by $\text{Cov}(\underline{S}, \underline{T}) = E[(\underline{S} - E(\underline{S}))(\underline{T} - E(\underline{T}))']$. For (random) vectors, the notation $||\underline{X}||_A^2 = \underline{X}' \underline{A}^{-1} \underline{X}$ is employed, where \underline{A} is a positive definite matrix.

Theorem 3.1. Let $\underline{Y}: q \times 1$, $\underline{X}: p \times 1$ be jointly distributed random vectors with $0 < \text{tr } \underline{\Sigma}_Y < \infty$; and let $\underline{A}: q \times q$ be positive definite. Then $\min E ||\underline{Y} - g(\underline{X})||_A^2$ occurs at $g(\underline{x}) = E(\underline{Y} | \underline{X} = \underline{x})$, where the minimum is over all measurable $g: R^p \rightarrow R^q$, for which $E ||\underline{Y} - g(\underline{X})||_A^2 < \infty$.

Proof. The result is immediate from the well-known identity

$$E ||\underline{Y} - g(\underline{X})||_A^2 = E ||\underline{Y} - E(\underline{Y} | \underline{X})||_A^2 + E ||g(\underline{X}) - E(\underline{Y} | \underline{X})||_A^2.$$

Theorem 3.2. Let $\underline{Y}: q \times 1$, $\underline{X}: p \times 1$ be jointly distributed random vectors with $0 < \text{tr } \underline{\Sigma}_Y < \infty$; and let $\underline{A}: q \times q$ be positive definite. Then $E ||\underline{Y} - E(\underline{Y} | \underline{X})||_A^2 = \text{tr}(\underline{A}^{-1} \underline{\Sigma}_Y) (1 - \{[\text{tr}(\underline{A}^{-1} \text{Cov } E(\underline{Y} | \underline{X}))] / [\text{tr}(\underline{A}^{-1} \underline{\Sigma}_Y)]\})$.

Proof. Note $E ||\underline{Y} - E(\underline{Y} | \underline{X})||_A^2 = E ||\underline{Y}||_A^2 - E ||E(\underline{Y} | \underline{X})||_A^2$. Without loss of generality, assume $E \underline{Y} = 0$. It is readily shown that

$E\|E(\underline{Y}|\underline{X})\|_{\underline{A}}^2 = \text{tr}[\underline{A}^{-1}\text{Cov}(E(\underline{Y}|\underline{X}))]$, and $E\|\underline{Y}\|_{\underline{A}}^2 = \text{tr}(\underline{A}^{-1}\underline{\Sigma}_Y)$, so that the result follows.

Observe that without any knowledge of \underline{X} the best expected minimum normed predictor of \underline{Y} is $E(\underline{Y})$ in the sense that $\min E\|\underline{Y}-\underline{c}\|_{\underline{A}}^2 = \text{tr}(\underline{A}^{-1}\underline{\Sigma}_Y)$ and occurs at $\underline{c} = E(\underline{Y})$. Thus, Theorem 3.2, like Theorem 2.1b, allows the measurement of the reduction in mean normed error due to knowledge of \underline{X} . This then suggests a suitable definition of a multivariate correlation ratio.

Definition 3.1. Let $\underline{Y}: q \times 1$ and \underline{X} be jointly distributed random vectors with $0 < \text{tr } \underline{\Sigma}_Y < \infty$. The correlation ratio of \underline{Y} upon \underline{X} is given by

$$\eta_{\underline{A}}(\underline{Y};\underline{X}) = \left(\frac{\text{tr}(\underline{A}^{-1}\text{Cov } E(\underline{Y}|\underline{X}))}{\text{tr}(\underline{A}^{-1}\underline{\Sigma}_Y)} \right)^{1/2},$$

where \underline{A} is a positive definite $q \times q$ matrix.

In the case when \underline{Y} is a scalar this definition clearly reduces to the previous definition (1.1) of correlation ratio. Obvious possible choices of \underline{A} are $\underline{A} = \underline{I}$ and $\underline{A} = \underline{\Sigma}_Y$, when $\underline{\Sigma}_Y$ is nonsingular. Note that $\eta_{\underline{I}}^2(\underline{Y};\underline{X}) = (\text{tr } \text{Cov } E(\underline{Y}|\underline{X})) / (\text{tr } \underline{\Sigma}_Y)$ and that $\eta_{\underline{\Sigma}_Y}^2(\underline{Y};\underline{X}) = q^{-1} \text{tr}(\underline{\Sigma}_Y^{-1/2} \text{Cov } E(\underline{Y}|\underline{X}) \underline{\Sigma}_Y^{-1/2})$. In practice, the actual choice of \underline{A} might depend on how difficult \underline{A} would be to estimate from the data for a particular multivariate model. It may be the case that $\text{tr } \underline{\Sigma}_Y$ is easier to estimate than the entire matrix $\underline{\Sigma}_Y$, so that $\underline{A} = \underline{I}$ should be chosen.

Recall that when \underline{Y} is a scalar, we have the important result that $\eta(\underline{Y};\underline{X}) = \max_{\underline{g}} \rho(\underline{Y};\underline{g}(\underline{X}))$. When \underline{Y} is a vector there is an analogous

result; however, we must introduce a suitable version of "correlation" between two vectors.

Definition 3.2. Let $\underline{S}: p \times 1$ and $\underline{T}: p \times 1$ be two jointly distributed random vectors with $0 < \text{tr } \underline{\Sigma}_S < \infty$, and $0 < \text{tr } \underline{\Sigma}_T < \infty$. The vector correlation between \underline{S} and \underline{T} is defined by

$$\rho_A(\underline{S}, \underline{T}) = \frac{\text{tr}(\underline{A}^{-1} \text{Cov}(\underline{S}, \underline{T}))}{(\text{tr}(\underline{A}^{-1} \underline{\Sigma}_S))^{1/2} (\text{tr}(\underline{A}^{-1} \underline{\Sigma}_T))^{1/2}},$$

where \underline{A} is a $p \times p$ positive definite matrix.

Note that $\rho_A(\underline{S}, \underline{T}) = \rho_A(\underline{T}, \underline{S})$, $|\rho_A(\underline{S}, \underline{T})| \leq 1$, and $\rho_A(\underline{S}, -\underline{T}) = -\rho_A(\underline{S}, \underline{T})$. For S, T scalars, $\rho_A(S, T) = \rho(S, T)$.

Theorem 3.3. Let $\underline{Y}: q \times 1$, $\underline{X}: p \times 1$ be jointly distributed random vectors with $0 < \text{tr } \underline{\Sigma}_Y < \infty$; and let $\underline{A}: q \times q$ be positive definite. Then $\max |\rho_A(\underline{Y}, g(\underline{X}))| = \eta_A(\underline{Y}; \underline{X})$, where the maximum is taken over all measurable functions $g: R^p \rightarrow R^q$, for which the vector correlation is defined.

Proof. Without loss of generality, assume $E(\underline{Y}) = \underline{0}$ and $Eg(\underline{X}) = \underline{0}$. From Theorem 3.1, it follows that

$$E \|\underline{Y} - E(\underline{Y}|\underline{X})\|_{\underline{A}}^2 \leq E \|\underline{Y} - c_g g(\underline{X})\|_{\underline{A}}^2, \quad (3.1)$$

for all g and constants c_g , which can depend on g . Expand (3.1) and choose $c_g = \{[\text{tr}(\underline{A}^{-1} \text{Cov } E(\underline{Y}|\underline{X}))]/[\text{tr}(\underline{A}^{-1} \text{Cov } g(\underline{X}))]\}^{1/2}$ to obtain

$$\text{Cov}(\underline{A}^{-1} \underline{Y}, E(\underline{Y}|\underline{X})) \geq c_g \text{Cov}(\underline{A}^{-1} \underline{Y}, g(\underline{X})). \quad (3.2)$$

Divide both sides of (3.2) by $[(\text{tr}(\underline{A}^{-1}\underline{\Sigma}_Y))(\text{tr}(\underline{A}^{-1}\text{Cov } E(\underline{Y}|\underline{X})))]^{\frac{1}{2}}$ to yield

$$\rho_A(\underline{Y}, E(\underline{Y}|\underline{X})) \geq \rho_A(\underline{Y}, g(\underline{X})), \quad (3.3)$$

for all measurable g . Because (3.3) holds for g and $-g$, and because $\rho_A(\underline{S}, -\underline{T}) = -\rho_A(\underline{S}, \underline{T})$, (3.3) holds with the right hand side of (3.3) in absolute values. Now

$$\begin{aligned} (\text{tr}(\underline{A}^{-1}\underline{\Sigma}_Y))^{\frac{1}{2}} \eta_A(\underline{Y}; \underline{X}) &= [\text{tr}(\underline{A}^{-1}\text{Cov } E(\underline{Y}|\underline{X}))]^{\frac{1}{2}} \\ &= [E\|\underline{E}(\underline{Y}|\underline{X})\|_A^2]^{\frac{1}{2}} \\ &= \frac{\text{tr}(\underline{A}^{-1}\text{Cov}(\underline{Y}; E(\underline{Y}|\underline{X})))}{[\text{tr}(\underline{A}^{-1}\text{Cov } E(\underline{Y}|\underline{X}))]^{\frac{1}{2}}} \end{aligned} \quad (3.4)$$

The result now follows from (3.3) and (3.4).

Corollary 3.1. Let $\underline{Y}: q \times 1$, $\underline{X}: p \times 1$ be jointly distributed random vectors with $0 < \text{tr } \underline{\Sigma}_Y < \infty$; and let $\underline{A}: q \times q$ be positive definite. Then $\eta_A(\underline{Y}; \underline{X}) = \rho_A(\underline{Y}, E(\underline{Y}|\underline{X}))$.

Proof. This is the result of (3.4).

Observe that if $E(\underline{Y}|\underline{X})$ is linear in \underline{X} , then $\eta_A(\underline{Y}; \underline{X}) = \max |\rho_A(\underline{Y}, \underline{B}\underline{X})|$, where the maximum is taken over all matrices $\underline{B}: q \times p$.

Before considering a number of examples, we briefly discuss further properties of the vector correlation coefficient when $\underline{A} = \underline{I}$. It is easily shown that

$$\rho_I(\underline{S}, \underline{T}) = \sum_{i=1}^p (\alpha_i \beta_i)^{1/2} \rho(S_i, T_i),$$

where $\underline{S} = (S_1, \dots, S_p)'$, $\underline{T} = (T_1, \dots, T_p)'$; and $\alpha_i = \text{Var } S_i / (\sum \text{Var } S_i)$, $\beta_i = \text{Var } T_i / (\sum \text{Var } T_i)$. Note $0 \leq \alpha_i, \beta_i \leq 1$ and $\sum \alpha_i = \sum \beta_i = 1$; thus, $\rho_I(\underline{S}, \underline{T})$ can almost be viewed as a weighted average of the pair-wise correlations. If $\text{Var } S_1 = \dots = \text{Var } S_p$ and $\text{Var } T_1 = \dots = \text{Var } T_p$, then $\rho_I(\underline{S}, \underline{T}) = (\sum_{i=1}^p \rho(S_i, T_i)) / p$. If $\text{Var } S_1 \gg \text{Var } S_2 = \dots = \text{Var } S_p$ and $\text{Var } T_1 \gg \text{Var } T_2 = \dots = \text{Var } T_p$, then $\rho_I(\underline{S}, \underline{T}) = \rho(S_1, T_1)$. Let $\underline{S}^* = \underline{\Gamma} \underline{S} + \underline{a}$, $\underline{T}^* = \underline{\Gamma} \underline{T} + \underline{b}$, where $\underline{\Gamma}$ is a $p \times p$ orthonormal matrix, and $\underline{a}, \underline{b}$ are arbitrary vectors; then $\rho_I(\underline{S}^*, \underline{T}^*) = \rho_I(\underline{S}, \underline{T})$.

Example 3.1. Multivariate Normal.

Let $(\underline{X}', \underline{Y}')' \sim N(\underline{0}, \underline{\Sigma})$, where \underline{X} is $q \times 1$, \underline{Y} is $p \times 1$ and $\underline{\Sigma}$ is partitioned similarly, i.e.,

$$\underline{\Sigma} = \begin{bmatrix} \underline{\Sigma}_X & \underline{\Sigma}_{XY} \\ \underline{\Sigma}_{XY}' & \underline{\Sigma}_Y \end{bmatrix}. \quad (3.5)$$

Then $E(\underline{X}|\underline{Y}) = \underline{\Sigma}_{XY} \underline{\Sigma}_Y^{-1} \underline{Y}$ and $\text{Cov}(E(\underline{X}|\underline{Y})) = \underline{\Sigma}_{XY} \underline{\Sigma}_Y^{-1} \underline{\Sigma}_{XY}'$; hence,

$$\eta_I(\underline{X}; \underline{Y}) = \{\text{tr}(\underline{\Sigma}_{XY} \underline{\Sigma}_Y^{-1} \underline{\Sigma}_{XY}')\}^{1/2} (\text{tr } \underline{\Sigma}_X)^{-1/2}$$

and

$$\eta_{\underline{\Sigma}_X}(\underline{X}; \underline{Y}) = (q^{-1} \text{tr}(\underline{\Sigma}_X^{-1} \underline{\Sigma}_{XY} \underline{\Sigma}_Y^{-1} \underline{\Sigma}_{XY}'))^{1/2}.$$

Observe that $\eta_{\underline{\Sigma}_X}(\underline{X}; \underline{Y}) = (q^{-1} \sum_{i=1}^q C_i^2(\underline{X}, \underline{Y}))^{1/2}$, where $C_i(\underline{X}, \underline{Y})$ is the i^{th} canonical correlation between \underline{X} , \underline{Y} , $i = 1, \dots, q$. If $p = q$, then

$$\rho_I(\underline{X}, \underline{Y}) = (\text{tr } \underline{\Sigma}_{XY}) \{ (\text{tr } \underline{\Sigma}_X) (\text{tr } \underline{\Sigma}_Y) \}^{-\frac{1}{2}} \quad \text{and} \quad \rho_{\underline{\Sigma}_X} = \text{tr}(\underline{\Sigma}_X^{-1} \underline{\Sigma}_{XY}) \{ p \text{tr}(\underline{\Sigma}_X^{-1} \underline{\Sigma}_Y) \}^{-\frac{1}{2}}.$$

Example 3.2. Elliptically Symmetric Distributions.

Let $\underline{Z} = (\underline{X}', \underline{Y}')'$ have p.d.f. $c|\underline{\Psi}|^{-\frac{1}{2}} \phi(\underline{z}'\underline{\Psi}^{-1}\underline{z})$. (Such distributions are called elliptically symmetric, e.g., Kelker (1970).) Let \underline{X} be $q \times 1$, \underline{Y} be $p \times 1$ and suppose $\underline{\Psi}$ is partitioned as in (3.5), where we assume $0 < \text{tr Cov}(\underline{X}) < \infty$. Then $E(\underline{X}|\underline{Y}) = \underline{\Psi}_{XY}\underline{\Psi}_Y^{-1}\underline{Y}$ and $\text{Cov}(E(\underline{X}|\underline{Y})) = \alpha \underline{\Psi}_{XY}\underline{\Psi}_Y^{-1}\underline{\Psi}_{XY}'$; also $\text{Cov}(\underline{X}) = \alpha \underline{\Psi}_X$, so that $\eta_I(\underline{X}; \underline{Y}) = \{ \text{tr}(\underline{\Psi}_{XY}\underline{\Psi}_Y^{-1}\underline{\Psi}_{XY}') \}^{\frac{1}{2}} (\text{tr } \underline{\Psi}_X)^{-\frac{1}{2}}$ and $\eta_{\underline{\Psi}_X}(\underline{X}; \underline{Y}) = \{ q^{-1} \text{tr}(\underline{\Psi}_X^{-\frac{1}{2}} \underline{\Psi}_{XY}\underline{\Psi}_Y^{-1}\underline{\Psi}_{XY}' \underline{\Psi}_X^{-\frac{1}{2}}) \}^{\frac{1}{2}}$.

Example 3.3. Multivariate Farlie-Gumbel-Morgenstern (FGM) Distribution.

In order to have tractable results, we consider a special four-dimensional FGM distribution with uniform marginals. (See Johnson and Kotz (1975) for the general multivariate distribution.) Let

$$\begin{aligned} f_{X_1, X_2, X_3, X_4}(x_1, x_2, x_3, x_4) &= 1 + \alpha_{12}v_{12} + \alpha_{13}v_{13} + \alpha_{14}v_{14} \\ &+ \alpha_{23}v_{23} + \alpha_{24}v_{24} + \alpha_{123}v_{123} \\ &+ \alpha_{124}v_{124} + \alpha_{1234}v_{1234}, \end{aligned} \quad (3.6)$$

where $v_{ij} = (1-2x_i)(1-2x_j)$, $v_{ijk} = (1-2x_i)(1-2x_j)(1-2x_k)$, and $v_{ijkl} = (1-2x_i)(1-2x_j)(1-2x_k)(1-2x_l)$. Direct calculation yields that $E(X_1|X_3, X_4) = (\frac{1}{2} - \alpha_{13}/6 - \alpha_{14}/6) + (\alpha_{13}/3)X_3 + (\alpha_{14}/3)X_4$, and $E(X_2|X_3, X_4) = (\frac{1}{2} - \alpha_{23}/6 - \alpha_{24}/6) + (\alpha_{23}/3)X_3 + (\alpha_{24}/3)X_4$. Further calculations yield that

$$\text{Cov}(E(X_1, X_2 | X_3, X_4))' = \begin{bmatrix} \frac{\alpha_{13}^2 + \alpha_{14}^2}{9} & \frac{\alpha_{13}\alpha_{23} + \alpha_{14}\alpha_{24}}{9} \\ \frac{\alpha_{13}\alpha_{23} + \alpha_{14}\alpha_{24}}{9} & \frac{\alpha_{23}^2 + \alpha_{24}^2}{9} \end{bmatrix} \quad (3.7)$$

and

$$\text{Cov}(X_1, X_2)' = \begin{bmatrix} 1/12 & \frac{\alpha_{12}}{36} \\ \frac{\alpha_{12}}{36} & 1/12 \end{bmatrix}. \quad (3.8)$$

Hence, for this example $\eta_I((X_1, X_2)'; (X_3, X_4)) = \{2/3(\alpha_{13}^2 + \alpha_{14}^2 + \alpha_{23}^2 + \alpha_{24}^2)\}^{1/2}$ and $\eta_{\Sigma_{12}}((X_1, X_2)'; (X_3, X_4)) = \{(2/9)(1 - 4\alpha_{12}^2)^{-1}[3(\alpha_{13}^2 + \alpha_{14}^2 + \alpha_{23}^2 + \alpha_{24}^2) - 2(\alpha_{12}\alpha_{13}\alpha_{23} + \alpha_{12}\alpha_{14}\alpha_{24})]\}^{1/2}$, where Σ_{12} is given by (3.8).

Suppose that f is given by (3.6), where $\alpha_{123} = \alpha_{124} = \alpha_{12} = 0$. By symmetry $\eta_I((X_3, X_4)'; (X_1, X_2)) = \{2/3(\alpha_{13}^2 + \alpha_{14}^2 + \alpha_{23}^2 + \alpha_{24}^2)\}^{1/2}$. Thus, if we set $\alpha_{13} = \alpha_{14} = \alpha_{23} = \alpha_{24} = 0$, we have $(X_1, X_2)'$ and $(X_3, X_4)'$ mutually have zero correlation ratio, each upon the other; yet $(X_1, X_2)'$ and $(X_2, X_4)'$ are not independent vectors. Very roughly speaking, α_{1234} measures a "higher order" level of dependence that is not measured in this example by η_I and $\eta_{\Sigma_{12}}$, where Σ_{12} is given by (3.8).

Example 3.4. Multinomial.

Let $\underline{X} = (X_1, \dots, X_I)'$ have p.m.f.

$$f_{\underline{X}}(\underline{x}) = \frac{n!}{\prod_{i=1}^I x_i! (n - \sum x_i)!} \prod_{i=1}^I p_i^{x_i} (1 - \sum p_i)^{n - \sum x_i}.$$

Then $E((X_1, \dots, X_J) | (X_{J+1}, \dots, X_I)) = n^*(p_1^*, \dots, p_J^*)$, where $n^* = n - (X_{J+1} + \dots + X_I)$.

and $p_i^* = p_i / (1 - (p_{J+1} + \dots + p_I))$, $i = 1, \dots, J$. It follows that $\text{Cov } E((X_1, \dots, X_J)' | (X_{J+1}, \dots, X_I)) = \Sigma_{1.2}$, where

$$\Sigma_{1.2} = \{n^*(p_i^* \delta_{ij} - p_i^* p_j^*)\}, \quad (3.9)$$

where $\delta_{ij} = 1$, if $i = j$, and $= 0$, otherwise. Since $\text{Cov}(X_1, \dots, X_J)' \equiv \Sigma_1$ is of the same form as (3.9), with n^* , p_i^* replaced by n , p_i , it follows that $\eta_1((X_1, \dots, X_J)'; (X_{J+1}, \dots, X_I)') = (n^*/n)(\Sigma p_i^*(1-p_i^*)) / (\Sigma p_i(1-p_i))$. Observe that $(\text{Cov}(X_1, \dots, X_J))^{-1} = \{n^{-1}(p_i^{-1} \delta_{ij} + (1 - \sum_{k=1}^J p_k)^{-1})\}$, so that $\eta_{\Sigma_1}((X_1, \dots, X_J)'; (X_{J+1}, \dots, X_I)') = (J)^{-1} [\sum_{i=1}^{J+1} ((1-p_i^*)(p_i^*/p_i))]$, where for $i = J+1$, $p_{J+1}^* = 1 - (p_1^* + \dots + p_J^*)$ and $p_{J+1} = 1 - (p_1 + \dots + p_J)$.

4. MOST PREDICTABLE LINEAR FUNCTIONS.

The multivariate correlation ratio $\eta_A(Y; X)$ measures the amount of predictability X has for Y for any jointly distributed random vectors X and Y . This relational notion is, of course, directional in that we are using the information in X in order to predict Y . To further understand this predictive relationship, it is natural to attempt to find what information in Y , that is, what function of Y , is most predictable from X . If we allow the consideration of all suitable measurable functions of Y , and follow the minimum norm approach with $A = I$, we are in essence, evaluating the sup-correlation $\rho'(X, Y)$ and finding the appropriate maximizing functions. As noted previously, this often is a difficult task mathematically. Consequently, we proceed analogously to the theory of principal components and canonical correlations, and restrict attention to linear functions of Y . The goal then is to find

which linear functions of Y are most predictable from X and to measure this degree of predictability.

For simplicity, we assume in the following that Σ_Y is non-singular; however, this could be avoided by examining determinantal equations.

Theorem 4.1. Let $Y: q \times 1$, $X: p \times 1$ be jointly distributed random vectors with $\text{tr } \Sigma_Y < \infty$ and Σ_Y nonsingular. Then

$$\max_{\beta} \eta^2(\beta'Y|X) = \lambda_{\text{MAX}}(\Sigma_Y^{-1/2} \text{Cov } E(Y|X) \Sigma_Y^{-1/2}), \quad (4.1)$$

where λ_{MAX} denotes the largest eigenvalue. Furthermore, this max occurs at $\beta = c \Sigma_Y^{-1/2} e_{\text{MAX}}$, where c is any nonzero constant, and e_{MAX} is an eigenvector corresponding to λ_{MAX} .

Proof.

$$\begin{aligned} \max_{\beta} \eta^2(\beta'Y|X) &= \max_{\beta} \frac{\text{Var } E(\beta'Y|X)}{\text{Var}(\beta'Y)} \\ &= \max_{\beta} \frac{\beta' \text{Cov } E(Y|X) \beta}{\beta' \Sigma_Y \beta} \\ &= \max_Y \frac{Y' \Sigma_Y^{-1/2} \text{Cov } E(Y|X) \Sigma_Y^{-1/2} Y}{Y' Y}, \end{aligned}$$

where $Y = \Sigma_Y^{-1/2} \beta$, so that the result is immediate (e.g., Rao (1973, p. 62)).

An immediate implication of Theorem 4.1 is the following corollary, whose proof is obvious.

Corollary 4.1. Let $\underline{Y}: q \times 1$, $\underline{X}: p \times 1$ be jointly distributed random vectors with $\text{tr } \Sigma_Y < \infty$ and Σ_Y nonsingular. Then

$$\begin{aligned} \text{Max}_{\substack{\beta: q \times 1 \\ g: R^p \rightarrow R^1}} \rho^2(\beta' \underline{Y}, g(\underline{X})) &= \lambda_{\text{MAX}}(\Sigma_Y^{-1/2} \text{Cov } E(\underline{Y}|\underline{X}) \Sigma_Y^{-1/2}), \end{aligned}$$

and occurs at $\beta = c \Sigma_Y^{-1/2} e_{\text{MAX}}$ and $g(x) = c e_{\text{MAX}}' \Sigma_Y^{-1/2} E(\underline{Y}|\underline{X})$.

The result of Corollary 4.1 is in some sense "half way" between linear canonical correlation for arbitrary random vectors and the sup-correlation of Sarmanov and Zaharov (1960).

For convenience, we introduce the following definition:

Definition 4.1. $L_1(\underline{Y}; \underline{X}) = \{\lambda_1(\Sigma_Y^{-1/2} \text{Cov } E(\underline{Y}|\underline{X}) \Sigma_Y^{-1/2})\}^{1/2}$, where λ_1 denotes the 1th largest eigenvalue.

When $\underline{Y}: q \times 1$ and $\underline{X}: p \times 1$ have a joint multivariate normal distribution $L_1(\underline{Y}; \underline{X}) = C_1(\underline{Y}, \underline{X})$, the first canonical correlation. For Example 3.3, $L_1^2((X_1, X_2)'; (X_3, X_4)')$ is the largest root of the determinantal equation $|\Sigma_{12 \cdot 34} - \lambda \Sigma_{12}| = 0$ where $\Sigma_{12 \cdot 34}$ is given by (3.7) and Σ_{12} by (3.8).

As in canonical correlation theory, multivariate versions of Theorem 4.1 could be obtained by examining uncorrelated iterations of (4.1). Another approach which we follow is to employ η_A . Suppose \underline{Y} is $q \times 1$ and \underline{X} is $p \times 1$; let B be an $r \times q$ matrix and A be an $r \times r$ positive definite matrix. Consider maximizing $\eta_A(B\underline{Y}; \underline{X})$ over B satisfying $\Sigma_{BY} = A^{-1}$. When $A = I$, this is equivalent to uncorrelatedness among the entries of $B \underline{Y}$.

Theorem 4.2. Let $\underline{Y}: q \times 1$, $\underline{X}: p \times 1$ be jointly distributed random vectors with $\text{tr } \underline{\Sigma}_Y < \infty$ and $\underline{\Sigma}_Y$ nonsingular; and let $\underline{A}: r \times r$ be positive definite. Then

$$\max_{\substack{\underline{B}: r \times q \\ \text{Cov}(\underline{B}\underline{Y}) = \underline{A}^{-1}}} \eta_A(\underline{B}\underline{Y}; \underline{X}) = \left\{ r^{-1} \sum_{i=1}^r L_1^2(\underline{Y}; \underline{X}) \right\}^{1/2}.$$

Proof. Observe that

$$\begin{aligned} \max_{\underline{B}} \eta_A^2 &= \max_{\underline{B}} \text{tr}(\underline{A}^{-1} \text{Cov } E(\underline{B}\underline{Y}|\underline{X})) / (\text{tr } \underline{A}^{-1} \underline{\Sigma}_{BY}) \\ &= \max_{\underline{B}} \frac{\text{tr}(\underline{A}^{-1/2} \underline{B} \text{Cov } E(\underline{Y}|\underline{X}) \underline{B}' \underline{A}^{-1/2})}{\text{tr}(\underline{A}^{-1/2} \underline{B} \underline{\Sigma}_Y \underline{B}' \underline{A}^{-1/2})} \\ &= r^{-1} \max_{\substack{\underline{C}: \underline{C}\underline{C}' = \underline{I}}} \text{tr } \underline{C} \underline{\Sigma}_Y^{-1/2} \text{Cov } E(\underline{Y}|\underline{X}) \underline{\Sigma}_Y^{-1/2} \underline{C}', \end{aligned}$$

where $\underline{C} = \underline{A}^{-1/2} \underline{B} \underline{\Sigma}_Y^{1/2}$. The result follows immediately (e.g., Rao (1973, pp. 63)).

For $\underline{X}, \underline{Y}$ having a multivariate normal distribution $\max_{\underline{B}} \eta_A^2(\underline{B}\underline{Y}; \underline{X}) = r^{-1} \sum_{i=1}^r C_i^2(\underline{Y}, \underline{X})$ i.e., the average of the squares of the first r canonical correlations.

An actual maximizing \underline{B} in Theorem 4.2 is given by

$$\underline{B} = \underline{A}^{1/2} (\underline{e}_1 : \dots : \underline{e}_r)' \underline{\Sigma}_Y^{-1/2},$$

where \underline{e}_i is an eigenvector corresponding to $L_i(\underline{Y}; \underline{X})$ and $\underline{e}_1, \dots, \underline{e}_r$ are orthogonal vectors.

The q values, $\{ \sum_{i=1}^r L_i^2(\underline{Y}; \underline{X}) \}^{1/2}$, $j = 1, \dots, q$, could themselves be viewed as measures of dependences of \underline{Y} upon \underline{X} and their properties explored. Note that knowing these q values is equivalent to knowing

the q eigenvalues, $L_i(\underline{Y}; \underline{X})$, $i = 1, \dots, q$.

For a fixed A , the estimation of $\eta_A(\underline{Y}; \underline{X})$ based upon n observations would be of interest, as would the estimation of $L_i(\underline{Y}; \underline{X})$. Both the actual estimation techniques employed and the resultant distribution theory would be dependent upon the underlying model assumptions.

ACKNOWLEDGMENT

The author wishes to thank Mr. Truc Nguyen for his help in Example 3.3.

REFERENCES

- Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis. Wiley, N.Y.
- Blakeman, J. (1905). "On tests of linearity of regression in frequency distributions". Biometrika 4, 332-350.
- Fisher, R. A. (1925). Statistical Methods for Research Workers. Oliver and Boyd, Ltd. 14th Edition (1970), Hafner Publishing Company.
- Fréchet, M. (1933). "Sur le coefficient, dit de corrélation". Comptes-Rendus de L'Académie des Sciences de Paris 197, 1268-1269.
- Fréchet, M. (1934). "Sur deux relations simples entre le 'coefficient' de corrélation et le 'rapport' de corrélation". Comptes-rendus du Deuxième Congrès des Mathématiciens des Pays Slavs, 209-210.
- Gebelein, H. (1941). "Das statistische problem der korrelations als variations und eigenwert problem und sein zusammenhang mit der ausgleichungsrechnung". Z. Angew. Math. Mech. 21, 364-379.
- Hall, W. J. (1970). "On characterizing dependence in joint distributions". In Essays in Probability and Statistics, Univ. of North Carolina Press, 339-376.
- Johnson, N. L. and Kotz, S. (1972). Distributions in Statistics: Continuous Multivariate Distributions. Wiley, N.Y.
- Johnson, N. L. and Kotz, S. (1975). "On some generalized Farlie-Gumbel-Morgenstern distributions". Communications in Statistics 4, 415-427.
- Kelker, D. (1970). "Distribution theory of spherical distributions and a location-scale parameter generalization". Sankya A 32, 419-430.
- Kotz, S. and Soong, C. (1977). "On measures of dependence: a survey of recent developments". Technical Report, Department of Mathematics, Temple University.
- Kruskal, W. H. (1958). "Ordinal measures of association". Journal of the American Statistical Association 53, 814-861.
- Lancaster, H. O. (1969). The Chi-Squared Distribution. Wiley, N.Y.
- Parzen, E. (1960). Modern Probability Theory and Its Applications. Wiley, N.Y.
- Pearson, K. (1903). "Mathematical contributions to the theory of evolution. On homotyposis in homologous but differentiated organs". Royal Society of London Proceeding 71, 288-313.

- Pearson, K. (1905). "Mathematical contributions to the theory of evolution XIV. On the general theory of skew correlation and non-linear regression". Drapers' Company Research Memoirs, Biometric Ser. II.
- Pearson, K. (1909). "On a new method of determining correlation between a measured character A and a character B, of which only the percentage of cases wherein B exceeds (or falls short of) a given intensity is recorded for each grade of A". Biometrika 7, 96-105.
- Pearson, K. (1948). Karl Pearson's Early Statistical Papers, ed. by E. S. Pearson. Cambridge University Press.
- Rao, C. R. (1973). Linear Statistical Inference and Its Applications 2nd Edition. John Wiley, N.Y.
- Rényi, A. (1959). "On measures of dependence". Acta. Math. Acad. Sci. Hungar. 10, 441-451.
- Sarmanov, O. V. (1958A). "The maximal correlation coefficient (symmetric case)". Dokl. Akad. Nauk. SSSR 120, 715-718. (English translation in Sel. Transl. Math. Statist. Probability 4, 271-275.)
- Sarmanov, O. V. (1958B). "The maximal correlation coefficient (non-symmetric case)". Dokl. Akad. Nauk. SSSR 121, 52-55. (English translation in Sel. Transl. Math. Statist. Probability 4, 207-210.)
- Sarmanov, O. V. and Zaharov, V. K. (1960). "Maximum coefficients of multiple correlation". Dokl. Akad. Nauk. SSSR 130, 269-271. (English translation in Soviet Math. 1, 51-53.)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|--|-----------------------|--|
| 1. REPORT NUMBER | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle) A Multivariate Correlation Ratio | | 5. TYPE OF REPORT & PERIOD COVERED Technical Report |
| 7. AUTHOR(s) Allan R. Sampson | | 6. PERFORMING ORG. REPORT NUMBER UP Stat. Report No. 80-6 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Pittsburgh Department of Mathematics & Statistics Pittsburgh, PA 15260 | | 8. CONTRACT OR GRANT NUMBER(s) F49620-79-C-0161 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) | | 12. REPORT DATE May 1980 |
| | | 13. NUMBER OF PAGES 19 |
| | | 15. SECURITY CLASS. (of this report) UNCLASSIFIED |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |
| 16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited. | | |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) | | |
| 18. SUPPLEMENTARY NOTES | | |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Correlation ratio, multivariate correlation ratio, vector correlation, canonical correlation, sup-correlation, elliptically symmetric. | | |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number) A brief review of the historical background and certain known results concerning the univariate correlation ratio are given. A multivariate correlation ratio of a random vector \underline{Y} upon a random vector \underline{X} is defined by $\eta_A(\underline{Y}; \underline{X}) = \text{tr}(\underline{A}^{-1} \text{Cov } E(\underline{Y} \underline{X}))^{1/2} (\text{tr}(\underline{A}^{-1} \underline{\Sigma}_Y))^{-1/2},$ | | |

where A is a given positive definite matrix. The properties of η_A are discussed, with particular attention paid to a "correlation-maximizing" property. A number of examples illustrating the application of η_A are given; these examples include the multivariate normal, the elliptically symmetric distributions, the Farlie-Morgenstern-Gumbel family, and the multinomial. The problem of maximizing $\eta_A(BY; X)$ over suitable matrices B is considered and the results that are obtained are related to canonical correlations for the multivariate normal.

END

DATE
FILMED

11-8

DTIC